DOCUMENT RESUME

ED 166 211 TM 007 809

AUTHOR Stallings, Jane A.
TITLE The Development of the Contextual Observation System.
INSTITUTION Stanford Research Inst., Menlo Park, Calif.
PUB DATE Mar 78
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Classroom Environment; *Classroom Observation Techniques; Data Processing; Early Childhood Education; Interpersonal Relationship; *Observation; *Test Construction; Test Reliability; Test Validity
IDENTIFIERS Interrater Reliability

ABSTRACT

The development of a contextual classroom observation system is described. The system consists of an instrument, trained observers, data handling, and analysis. Thus, instrument development was necessarily paralleled by the development of training procedures, data recording and processing systems, programming, and analytic techniques. The comprehensive observation system described was initially developed for use in the evaluation of Head Start and Follow Through programs, but has received wide application and has led to the development of a number of derivative observation systems. Data collection forms concerning the physical classroom environment, the snapshot of the classroom, and five-minute samples of interpersonal relations are included. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED166211

SRI International

IM002 809

Session 24.22
Division C

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE DEVELOPMENT OF THE CONTEXTUAL OBSERVATION SYSTEM

Jane A. Stallings, Ph.D.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Jane Stallings

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Presented to:

American Educational Research Association
Toronto, Canada
March 1978

SRI International

333 Ravenswood Ave. • Menlo Park, California 94025
(415) 326-6200 • Cable: STANRES, Menlo Park • TWX: 910-373-1246

THE DEVELOPMENT OF THE CONTEXTUAL OBSERVATION SYSTEM

The purpose of this paper is to describe how a contextual observation system developed. The system described consists of an instrument, trained observers, data handling, and analysis. Thus, instrument development was necessarily paralleled by the development of training procedures, data recording and processing systems, programming, and analytic techniques. The comprehensive observation system described was initially developed for use in the evaluation of Head Start and Follow Through programs, but has received wide application and has led to the development of a number of derivative observation systems.

Development

The observation system used at SRI International was developed in response to a request by the U.S. Office of Education to evaluate the implementation of the models of education found in Follow Through Planned Variation. These models vary from highly structured, teacher directed models to flexible open-education models. Of interest to the government was whether or not the model specifications were being followed: whether the correct materials were being provided and used, whether specified activities of children, teachers, and aides were occurring with the specified frequency, and whether the interaction patterns of all the participants followed sponsor guidelines. This type of systematic documentation required a comprehensive observation system. As we studied the available systems, it became clear that while one system documented interactions and another listed materials or child behaviors, no existing system provided the complete picture of the classroom that was required.

A record to describe the physical environment was not difficult to develop; it emerged from a list of descriptors each Follow Through model sponsor prepared of the materials that should be used in the classroom.

This long list indicated whether specific materials were present and used (see Figure 1).

The instrument that was developed to record the classroom activities was more complicated. Not only did we want to know what activities occurred throughout the day, we also wanted to know what activities occurred together. Some models indicated that reading, math, and art should be scheduled for specific periods during the day, but other models indicated that reading, math, and art should occur simultaneously, with groups of children engaged in the different activities. Also of interest was which adult was working with the children in the activities, or whether the children were working independently. To capture this type of information, we developed a grid that listed all the possible classroom activities (see Figure 2) and the possible grouping arrangements. This allowed us to examine who the teacher and aides worked with and in what activities. Questions could be answered such as does the teacher or aide work with one child, two children, small groups, or large groups in reading, math, or art? Also, we could compute the average frequency with which students were involved in reading or any other activity throughout the day. We have come to call this grid the Snapshot, since it provides a still picture of everyone in the classroom. The grid was filled out four times an hour, 5 hours a day, for 3 days, for a total of 60 Snapshots for each classroom. A very good description of what goes on, showing teacher, aide, and child activities is provided. Also, the activities of Day 1 can be compared with the activities of Day 2 and Day 3.

Once we were able to describe the classroom stage and the scene, we wanted to document the interaction of the actors--teachers, aides, and children. The Flanders system provides a good model for recording these interactions. We were fortunate that Dr. Ned Flanders was near at hand at the Stanford University Center for Advanced Study in the Behavioral Sciences. He generously gave of his time and assisted us in structuring our interaction system. It still contains many of his codes, but also includes codes for nonverbal actions, movements, and many descriptors of the message (see Figure 3). The interactions of a teacher, an aide, or a child were recorded for a 5-minute period, four times an hour, 5 hours

Physical Environment Information
(Mark all that apply.)

For each of the items below, mark all that apply:

① **Present**
② **Used today**

Seating Patterns:
- ○ Movable tables and chairs for seating purposes.
- ○ Stationary desks in rows.
- ○ Assigned seating for at least part of the day.
- ○ Children select their own seating locations.
- ○ Teacher assigns children to groups.
- ○ Children select their own work groups.

GAMES, TOYS, PLAY EQUIPMENT
① ② small toys (trucks, cars, dolls and accessories
① ② puzzles, games
① ② wheel toys
① ② small play equipment (jumpropes, balls)
① ② large play equipment (swings, jungle gym)
① ② children's storybooks
① ② animals, other nature objects
① ② sandbox, water table
① ② carpentry materials, large blocks
① ② cooking and sewing supplies

INSTRUCTIONAL MATERIALS
① ② Montessori, other educational toys
① ② children's texts, workbooks
① ② math/science equipment, concrete objects
① ② instructional charts

AUDIO, VISUAL EQUIPMENT
① ② television
① ② record or tape player
① ② audio-visual equipment

GENERAL EQUIPMENT, MATERIALS
① ② children's own products on display
① ② displays reflecting children's ethnicity
① ② other displays especially for children
① ② magazines
① ② achievement charts
① ② child-size sink
① ② child-size table and chairs
① ② child-size shelves
① ② arts and crafts materials
① ② blackboard, feltboard
① ② child's own storage space
① ② photographs of the children on display

OTHER
① ② please specify . . . . . .

MAKE NO STRAY MARKS IN BLANK AREAS

FIGURE 1 PHYSICAL ENVIRONMENT

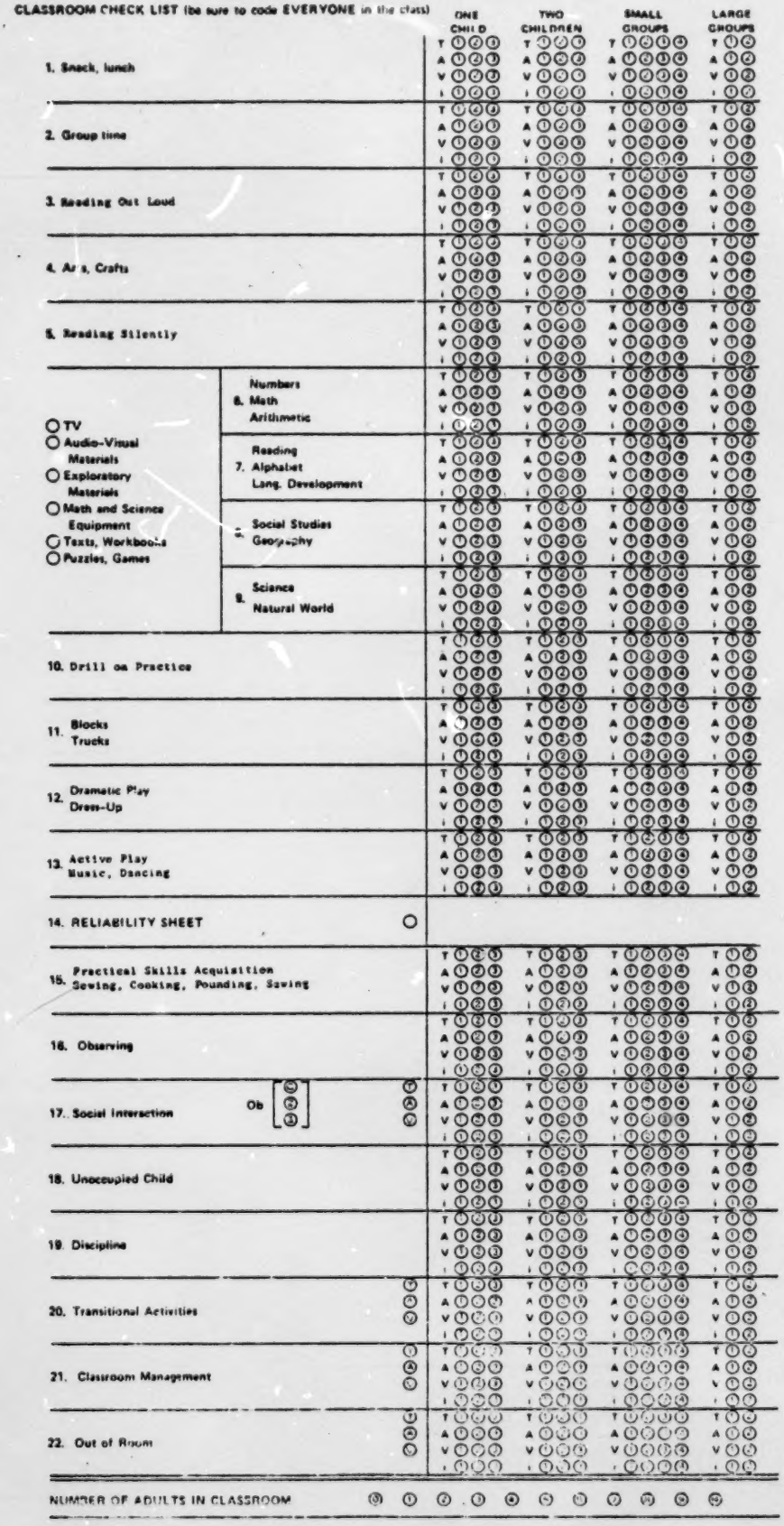
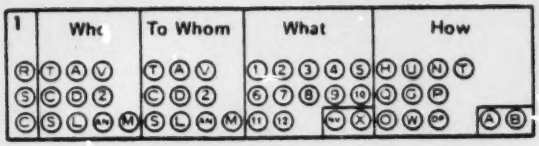


FIGURE 2 THE SNAPSHOT



Who/To Whom	What	How
T - Teacher	1 - Command or Request	H - Happy
A - Aide	2 - Open-ended Question	U - Unhappy
V - Volunteer	3 - Response	N - Negative
C - Child	4 - Instruction, Explanation	T - Touch
D - Different Child	5 - Comments, Greetings; General Action	Q - Question
2 - Two Children	6 - Task-related Statement	G - Guide/Reason
S - Small Group (3-8)	7 - Acknowledge	P - Punish
L - Large Group (9 up)	8 - Praise	O - Object
An - Animal	9 - Corrective Feedback	W - Worth
M - Machine	10 - No Response	DP - Dramatic Play/Pretend
	11 - Waiting	A - Academic
	12 - Observing, Listening	B - Behavior
	NV - Nonverbal	
	X - Movement	

R - Repeat the frame
S - Simultaneous action
C - Cancel the frame

FIGURE 3 LIST OF CODES FOR FIVE-MINUTE INTERACTION

a day. For the 3 days of observation, 60 interaction observations were made for each classroom, in addition to the 60 Snapshots.

The physical environment checklist, the snapshot of activities and groupings, and the 5-minute interactions observation (FMI) together constitute a comprehensive observation system that describes the total classroom environment.

Before the first observation system was put into final form, the principal researcher tried it out in the classrooms of eight different models. At the same time, ethnographic recordings were made of the classrooms so that comparisons could be made between the written script and the coded data. Several important occurrences found only in the written script, such as nonverbal feedback through touching, were then incorporated into the coding system.

Fortunately, we had the opportunity to observe in the Follow Through classrooms over a period of 4 years. The observation system was modified over time. Each year, representatives of the Follow Through models were asked to critique the system and recommend changes that would more specifically describe their classrooms. By the fourth year, most of the activities and behaviors that could be seen or heard in classrooms could be recorded on this observation system. Some limitations of the system are that the more subtle objectives of educational models, such as motivating children through intrinsic rewards, cannot be recorded or quantified.

Training

Proper use of the observation system required intensive observer training. The first series of training sessions were conducted by the principal researcher and two assistants. Observer trainees were brought from their communities to a regional center for 7 days of training. Observers had already received home training packets and were expected to have memorized the code definitions before arriving at the sessions. We tried to select trainees who had at least 2 years of college, and who had some experience with young children. We exempted those people who had any association with the Follow Through program. This exemption rule was

initiated after one observer who was a Follow Through parent told the teachers in the New York school where she was observing that the model in her child's school was much better. This statement caused concern about the objectivity of the study. In another case in Arizona, an observer who was the wife of a principal in a Follow Through school was advising the teachers about how to carry out the model. Thus, training covered not only memorizing the codes and developing the concepts, but also appropriate observer behavior within the community.

During training, for the most part, we worked in small groups, trying to keep the ratio of trainees to trainer at a 5 to 1 ratio. We used many training techniques, including role playing, coding videotapes, coding natural classrooms, and completing homework assignments where the trainees wrote vignettes using the codes. At some time, each observer coded in the classroom with an SRI trainer. At the completion of training, to help decide whether the observers were competent to collect data, we gave a criterion videotape test which showed examples of all of the codes. Any observer who could not code each code with at least 70% reliability was not permitted to collect data. Those who passed the test were put to work immediately so that the skill they had attained would not be lost from lack of use.

Assessing Reliability

During the first 2 years, interrater agreement was assessed in the classrooms using SRI trainers as the criterion. Sending a person to 24 locations all across the country proved to be a very expensive process. What's more, we could not control the number or variety of codes that would occur naturally in the classroom during the hours when the two raters were recording together. Sometimes as few as 15 out of 42 codes were used. This left us uncertain about how the observer might code the remaining 27 codes. During the third and fourth years, the ability of the observers to code correctly was assessed using the criterion videotape test only. To see whether this ability improved or deteriorated, the videotape was administered 2 weeks after training and data gathering had started. In the 7 years of our experience in training observers, we have always found

that most observers gain rather than lose in their ability to code if they continue to use the system.

Interrater agreement in classrooms was compared with scores on the criterion videotapes in a subsample of classrooms. We found that the two methods of assessing observer ability to code reliably provided similar results. Observers who coded a particular code incorrectly on the videotape tended to disagree with the SRI trainer when that code occurred during the interrater agreement sessions. Since that time we have tended to use the criterion videotapes to judge quality of observer coding, since this method is less expensive and ensures examining a wide range of codes.

Validity

The validity of the instrument has been examined in several ways. First, the face validity has been carefully studied by the Follow Through models. The operational definitions are very straightforward: to code "positive affect," an observer must see a broad smile or hear laughter; to code punishment, the observer must see the teacher restrain or touch a child and/or hear them withhold some privilege. The code name reflects a concept whose implementation can be seen or heard. (See the operational definitions provided in Stallings and Kaskowitz, 1974.)

Cross-validity checks have been made comparing data from the Snapshot instrument with the data from the interaction instrument. For example, the proportion of time for teachers working with one child recorded on the Snapshot is similar to the proportion of time teachers speak to small groups on the FMI observation, and these variables also show a similar correlation with outcome scores.

Another indicator of validity is that the frequency of occurrences of specified behaviors in the Follow Through classrooms were similar to what would be predicted given the descriptions and stated goals of the models. Using the observation data, 72% of 271 classrooms were correctly identified as to sponsor names.

Yet another indicator of validity is that in the overall study of 58 classrooms relating classroom processes to child outcomes, those models that more frequently had children asking questions (a variable related to nonverbal problem solving) also had children scoring higher on the nonverbal problem solving test.

Data Processing

Our data processing techniques have improved tremendously during the past 7 years. The first year observers entered the data in booklets using alpha and numeric digits. The data were handscored for the 24 classrooms in the sample and keypunched; the process took all summer. When no praise was entered on the computer record for the University of Kansas--a behavior modification model--I knew something was wrong, so we found the flaw in the program and started over again.

The second year we changed to an optical scan format. This required observers to mark the form rather than write the codes. While there were some problems in marking correctly, we found that Clorox on a cotton swab could remove a wrong mark.

The scanning system was much more efficient and we could collect more discrete kinds of information, such as a second language being spoken or objects being used in instruction. Very complicated computer programs were required to translate the data from the optical scan form to a usable tape of frequency counts of each code per classroom. Over the past 7 years these programs have been refined, and today our turn-around time for processing data has been reduced from 6 months to 2 weeks.

The programs used to develop the lists of variables from the codes have also emerged over time. Programming these observation data was a new experience for SRI personnel. The programmers were creative but only 24 variables could be developed during the time period allowed. The second year, the optical scan form required a different program format, and we programmed 13 Snapshot variables and 53 FMO interaction variables. We wanted to count particular sequences of variables such as (1) teacher asks a small group a simple direct question, followed by (2) the small

group responding, followed by (3) the teacher giving some feedback, but that kind of programming had to be foregone until the third year. In the third year the programming was difficult but successful. We had 52 Snapshot variables, 120 FMO interaction variables, and four sequence variables. Unfortunately, the documentation of the programs was not completed before the chief programmer left. Another gifted programmer took over in the fourth year and developed a very simple format that could be easily documented. This system is now very efficient and can easily be adapted to any of our observation instrument modifications.

We can now use a mega variable such as how frequently does any adult work with one child or a metavariable such as how frequently does a teacher or an aide or a volunteer work with one child in reading. Also, twelve different types of feedback can now be documented using a combination of ten codes in different patterns.

Instrument Modification

The system carefully developed for the Follow Through evaluation has since been modified for use in three Teacher Corps studies, two high school studies, three preschool studies, two Canadian studies (with a French translation), and two studies of the California Early Childhood Education (ECE) program. This wide use shows that operational definitions can be changed but the system works and stays the same. Our California ECE study was interesting in that it examined teaching processes of 48 third grade teachers and related those to the children's gains in reading. Although this study took place four years after the last Follow Through observation study, and the sample contained only California classrooms rather than a national sample, the findings were very similar to those of Follow Through. Direct teaching methods were the most efficient in teaching reading. An analysis of variance of 14 classrooms matched for entering third grade test scores showed clearly what was occurring in the seven classrooms that gained compared to the seven classrooms that did not gain. The findings are so specific that inservice teacher training workshops can be developed to instruct teachers to increase certain behaviors and decrease other behaviors.

Summary

Observation technology has come a long way since 1969. We are grateful to Ned Flanders, Don Medley, and others who pioneered this field. Findings have become more and more specific and useful. We are building a body of knowledge. Certainly the way teachers and students are being evaluated is changing and we are pleased to be part of the process.